Generalized Variable Step Size Continuous Mixed *p*-Norm Adaptive Filtering Algorithm

Long Shi, Haiquan Zhao, *Senior Member, IEEE*, and Yuriy Zakharov, *Senior Member, IEEE*

***Abstract*— To further improve the performance of the variable step size continuous mixed *p*-norm (VSS-CMPN) adaptive filtering algorithm in the presence of impulsive noise, a generalized VSS-CMPN algorithm (GVSS-CMPN) is proposed in this paper. Instead of assuming the probability density-like function $\lambda(p)$ to be uniform, a linear function is proposed for $\lambda(p)$ to control the mixture of various error norms. The influence of the selection of the regulating factor (slope of the linear function) is discussed. Besides, the computational complexity as well as the mean-square convergence analysis is presented. Simulations conducted in the system identification scenario demonstrate the superiority of the proposed algorithm over known algorithms.**

***Index Terms*—impulsive noise, mean-square convergence, mixed *p*-norm, probability density-like function**

I. Introduction

Over the past few decades, the least-mean-squares (LMS) algorithm has attracted much attention due to its low computational cost and easy implementation [1]. Unfortunately, impulsive noise which is frequently encountered in practice would deteriorate the performance of the LMS algorithm since it is based on the mean square error criterion [2]-[4].

To address this problem, the least absolute deviation (LAD) algorithm and the robust mixed-norm (RMN) algorithm were proposed [5]-[7]. The later algorithm is based upon the convex combination of the LMS and LAD algorithms, and employs a scalar parameter considered as the probability that the instantaneous desired response does not contain significant impulsive noise to control the mixture [7]. Subsequently, its normalized version (NRMN) was developed in [8], which provides an improved performance in the non-stationary environment. Recently, the continuous mixed *p*-norm (CMPN) adaptive filtering algorithm was proposed [9], which combines various *p*-norms for $1 \le p \le 2$ using a continuous probability density density-like function $\lambda(p)$. Based on different approximations for the expectation of the l_p-norm of the error signal, the variable step size CMPN (VSS-CMPN) algorithm and the block CMPN (Block-CMPN) algorithm were put forward [9]. The VSS-CMPN algorithm relies on single point estimate, while the Block-CMPN algorithm relies on a block of the error signal. As reported in [9], the VSS-CMPN algorithm outperforms the Block-CMPN algorithm because the latter is derived from past error samples. Based on the *p*-norm error criterion, several algorithms have been investigated [10]-[13].

However, in the CMPN, the probability density-like function $\lambda(p)$ assumed to be uniform, i.e., $\lambda(p)=1$ for $p \in [1,2]$, is just a special case, and other situations that meet the constraint $\int_1^2 \lambda(p)\mathrm{d}p = 1$ have not been taken into account. Therefore, in this brief, we focus on further development of the VSS-CMPN algorithm. As can be seen in Fig. 1, rotating the function $\lambda(p)$ around the coordinate $\left(\frac{3}{2}, 1\right)$ does not change the value of the integral $\int_1^2 \lambda(p)\mathrm{d}p$, which implies that the rotated function can be utilized as $\lambda(p)$. The rotated function can be expressed as a linear function

$$\lambda(p) = \theta\left(p - \frac{3}{2}\right) + 1 \tag{1}$$

where θ is a constant denoting the regulating factor. Although, strictly speaking $\lambda(p)$ is not a probability density function (PDF), we still treat it as a PDF and correspondingly limit the range of θ. Since the PDF $\lambda(p)$ cannot be negative, θ belongs to the interval $[-2, 2]$. Based on (1), we develop a generalized VSS-CMPN (GVSS-CMPN) algorithm in this paper. We discuss the influence of the regulating factor θ employing a nonlinear function of the error signal and reveal the connection between the regulating factor and the algorithm robustness. We also compare the computational complexity of the proposed GVSS-CMPN algorithm with that of several existing algorithms. In addition, we provide the mean-square convergence analysis. Simulations conducted in the system identification scenario illustrate the advantages of our finding. In general, our main contributions are threefold:

(a) A linear function is put forward for the PDF $\lambda(p)$ to derive the GVSS-CMPN algorithm.

(b) The influence of the selection of the regulating factor is discussed by using a nonlinear function of the error signal.

(c) The mean-square convergence analysis for the proposed algorithm is presented.

This work was partially supported by National Science Foundation of P.R. China (Grant: 61871461, 61571374 and 61433011). The work of Y. Zakharov was partly supported by the U.K. EPSRC through Grants EP/P017975/1 and EP/R003297/1.

Long Shi and Haiquan Zhao are with Key Laboratory of Magnetic Suspension Technology and Maglev Vehicle, Ministry of Education, and also with School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, People's Republic of China (e-mail: lshi@my.swjtu.edu.cn; hqzhao_swjtu@126.com).

Yuriy Zakharov is with Department of Electronic Engineering, University of York, York YO10 5DD, U.K. (yury.zakharov@york.ac.uk)

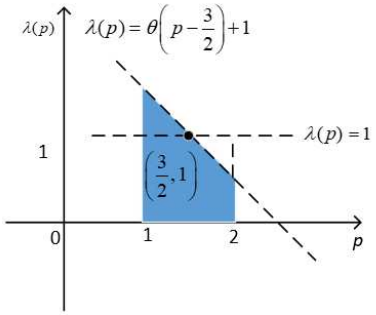

Fig. 1. The linear function designed for the PDF $\lambda(p)$

II. Proposed Algorithm

Consider an unknown M-dimensional vector $\mathbf{w}_o$ that satisfies the following linear model

$$d(k) = \mathbf{u}^T(k)\mathbf{w}_o + \eta(k) \tag{2}$$

where $d(k)$ denotes the desired signal, $\mathbf{u}(k) = [u(k), u(k-1), \cdots, u(k-M+1)]^T$ stands for the input vector, $\eta(k)$ is the zero-mean background noise, and $(\cdot)^T$ represents vector or matrix transpose.

Inspired by the CMPN algorithm [9], the cost function of the proposed scheme is given by

$$J(k) = \int_1^2 \lambda(p) E\left\{|e(k)|^p\right\} \mathrm{d}p \tag{3}$$

where $e(k) = d(k) - \mathbf{u}^T(k)\mathbf{w}(k)$ is the error signal, and $E\{\cdot\}$ denotes the expectation.

The proposed algorithm is obtained by applying the steepest descent principle

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \mu \nabla_{\mathbf{w}(k)} J(k) \tag{4}$$

where μ denotes the fixed step size ($\mu > 0$), and $\nabla_{\mathbf{w}(k)} J(k)$ is given by

$$\nabla_{\mathbf{w}(k)} J(k) = \int_1^2 \lambda(p) \frac{\partial}{\partial \mathbf{w}(k)} E\left\{|e(k)|^p\right\} \mathrm{d}p \tag{5}$$

Employing the approximation method of single point estimate and replacing $E\left\{|e(k)|^p\right\}$ with $|e(k)|^p$, we obtain

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \gamma_k \operatorname{sign}(e(k))\mathbf{u}(k) \tag{6}$$

where $\gamma_k = \int_1^2 p\lambda(p)|e(k)|^{p-1} \mathrm{d}p$, and sign$(\cdot)$ stands for the sign function.

With (1), γ_k can be rewritten as

$$\begin{aligned} \gamma_k &= \int_1^2 p\left[\theta\left(p - \frac{3}{2}\right) + 1\right]|e(k)|^{p-1} \mathrm{d}p \\ &= \int_1^2 p\theta\left(p - \frac{3}{2}\right)|e(k)|^{p-1} \mathrm{d}p + \int_1^2 p|e(k)|^{p-1} \mathrm{d}p \\ &= \theta\delta(k) + \rho(k) \end{aligned} \tag{7}$$

where

$$\delta(k) \triangleq \int_1^2 p\left(p - \frac{3}{2}\right)|e(k)|^{p-1} \mathrm{d}p \tag{8}$$

and

$$\rho(k) \triangleq \int_1^2 p|e(k)|^{p-1} \mathrm{d}p \tag{9}$$

Finally, $\delta(k)$ and $\rho(k)$ are expressed as

$$\delta(k) = \frac{2|e(k)| - 2}{\ln^3|e(k)|} + \frac{0.5 - 2.5|e(k)|}{\ln^2|e(k)|} + \frac{|e(k)| + 0.5}{\ln|e(k)|} \tag{10}$$

$$\rho(k) = \frac{\ln|e(k)|(2|e(k)| - 1) - |e(k)| + 1}{\ln^2|e(k)|} \tag{11}$$

Note that, when $\theta = 0$, the GVSS-CMPN algorithm reduces to the VSS-CMPN algorithm, which implies that the GVSS-CMPN algorithm is a generalized version of the VSS-CMPN algorithm.

A. Selection of the Regulating Factor

In order to investigate the influence of the regulating factor θ on the algorithm performance, we seek to explore this issue using a nonlinear function of the error signal. Before proceeding further, (6) is reformulated as

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu f(e(k))e(k)\mathbf{u}(k) \tag{12}$$

where $f(e(k)) = \dfrac{\gamma_k}{|e(k)|}$ is a nonlinear function of the error signal, and $\mu f(e(k))$ can be viewed as the overall step size for the LMS algorithm. Using (7), $f(e(k))$ can be further written as

$$f(e(k)) = \frac{\theta\delta(k) + \rho(k)}{|e(k)|} \tag{13}$$

Fig. 2 plots the nonlinear function $f(e(k))$ with different regulating factors. As can be seen, the nonlinear function $f(e(k))$ with $\theta = 2$ yields a larger value than that with other tested values under the same $|e(k)|$. With the decrease of θ, $f(e(k))$ is reduced for the same $|e(k)|$.

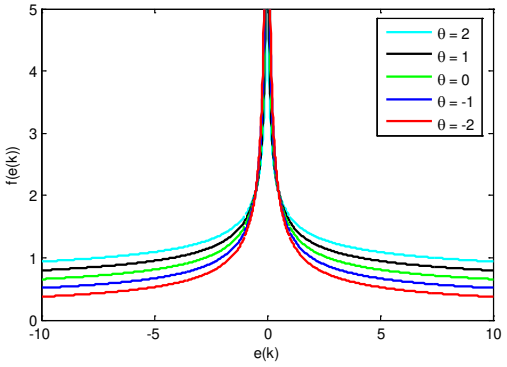

Fig. 2. The nonlinear function $f(e(k))$ with different θ

When the impulsive noise occurs, the absolute value of the error signal, i.e., $|e(k)|$, is very large. We expect the step size to be reduced to suppress the impulsive noise, ensuring the robustness of the algorithm. Since μ is fixed, $f(e(k))$ should be a small value to achieve this goal. Therefore, according to the characteristics of the function described previously, adopting a small regulating factor θ is beneficial to improve the robustness of the algorithm. Moreover, the closer the

regulating factor θ is to -2, the more robust against impulsive noise the algorithm is. When there is no impulsive noise, the error signal does not exhibit great value, and we hope to speed up the convergence by increasing the step size. Therefore, θ should be close to 2 to obtain a large value for $f(e(k))$.

B. Practical Consideration

It is observed from Fig. 2 that as $|e(k)|$ tends to zero, the function $f(e(k))$ tends to infinity. When the algorithm arrives at the steady state, if $|e(k)|$ is small enough, $f(e(k))$ will be high, which may result in instability of the algorithm. To prevent this happening, we set a threshold for $|e(k)|$ as

$$|e(k)| = \begin{cases} \kappa\sigma_\eta, & \text{if } |e(k)| \le \kappa\sigma_\eta \\ |e(k)|, & \text{otherwise} \end{cases} \tag{14}$$

where κ is a positive constant, and σ_η denotes the standard deviation of the background noise. The threshold $\kappa\sigma_\eta$ is motivated by the fact that $E\left[e^2(\infty)\right] \approx \sigma_\eta^2$ [1]. Replacing $E\left[e^2(\infty)\right]$ with $e^2(\infty)$ and performing a square root operation, we have $|e(\infty)| \approx \sigma_\eta$. We then introduce an adjustment parameter κ to generate the threshold $\kappa\sigma_\eta$. If the variance σ_η^2 of the background noise is unknown, the online learning method can be used to estimate it [14]. Extensive simulations suggest that $\kappa \le 1$ can provide good performance.

C. Computational Complexity

Table I compares the computational complexity of various algorithms, where N_w denotes the length of sliding window of the RMN algorithm. It is observed that the RMN algorithm is more computationally expensive than the LAD algorithm because the calculation for mixing parameter in the RMN increases the cost. As compared to the VSS-CMPN algorithm, the proposed GVSS-CMPN algorithm requires more multiplications, additions and logarithmic operations since the additional computation in (10) and (11) is needed. However, this increase is moderate.

TABLE I
COMPUTATIONAL COMPLEXITY OF VARIOUS ALGORITHMS

Algorithms	Multiplications	Additions	Logarithmic operations
LAD [4]	$2M+1$	$2M-1$	0
RMN [6]	$2M+2N_w+4$	$2M+N_w-1$	0
VSS-CMPN [8]	$2M+4$	$2M$	2
GVSS-CMPN	$2M+10$	$2M+2$	5

III. MEAN-SQUARE STABILITY

To perform the convergence analysis in the mean-square sense, we define the weight error vector $\tilde{\mathbf{w}}(k)$ as

$$\tilde{\mathbf{w}}(k) = \mathbf{w}_o - \mathbf{w}(k) \tag{15}$$

Combining (15) with (12) gives rise to

$$\tilde{\mathbf{w}}(k+1) = \tilde{\mathbf{w}}(k) - \mu f(e(k))e(k)\mathbf{u}(k) \tag{16}$$

Taking the square and expectation of both sides of (16) yields

$$\begin{aligned} E\left[\|\tilde{\mathbf{w}}(k+1)\|^2\right] &= E\left[\|\tilde{\mathbf{w}}(k)\|^2\right] - 2\mu E\left[f(e(k))e(k)e_a(k)\right] \\ &\quad + \mu^2 E\left[f^2(e(k))e^2(k)\|\mathbf{u}(k)\|^2\right] \\ &= E\left[\|\tilde{\mathbf{w}}(k)\|^2\right] - \Delta(k) \end{aligned} \tag{17}$$

where $e_a(k) = \tilde{\mathbf{w}}^T(k)\mathbf{u}(k)$ denotes noise-free a *priori* error signal, and $\Delta(k)$ is given by

$$\Delta(k) = 2\mu E\left[f(e(k))e(k)e_a(k)\right] - \mu^2 E\left[f^2(e(k))e^2(k)\|\mathbf{u}(k)\|^2\right] \tag{18}$$

From (17), the convergence of the GVSS-CMPN algorithm in the mean-square sense can be guaranteed if the squared weight error satisfies

$$E\left[\|\tilde{\mathbf{w}}(k+1)\|^2\right] \le E\left[\|\tilde{\mathbf{w}}(k)\|^2\right] \tag{19}$$

The inequality (19) is established as long as $\Delta(k) \ge 0$. Therefore, a mean-square convergence condition is achieved for

$$0 < \mu < \mu_m = \frac{2E\left[f(e(k))e(k)e_a(k)\right]}{E\left[f^2(e(k))e^2(k)\|\mathbf{u}(k)\|^2\right]} \tag{20}$$

where μ_m denotes the upper bound of the step size.

To proceed, we make the following assumptions:

A1: The background noise $\eta(k)$ is independent of the input signal $\mathbf{u}(k)$ and noise-free a *priori* error signal $e_a(k)$ [1].

A2: The ratio of the expectation of two random variables is approximated by the expectation of the ratio between them, i.e., $\frac{E(x)}{E(y)} \approx E\left\{\frac{x}{y}\right\}$, which is reasonable for sufficiently long filters [15].

A3: The error signal $e(k)$ is uncorrelated with the input signal $\mathbf{u}(k)$ [16].

Combining $e(k) = e_a(k) + \eta(k)$ and applying Assumption **A1** for μ_m in (20), we arrive at

$$\mu_m = \frac{2E\left[f(e(k))e_a^2(k)\right]}{E\left[f^2(e(k))e^2(k)\|\mathbf{u}(k)\|^2\right]} \tag{21}$$

Define the auto-correlation matrix as $\mathbf{R}_{\mathbf{uu}} = E\left[\mathbf{u}(k)\mathbf{u}^T(k)\right]$. Invoking Assumptions **A2** and **A3** for (21), we obtain

$$\mu_m = 2E\left\{\frac{e_a^2(k)}{f(e(k))e^2(k)\|\mathbf{u}(k)\|^2}\right\} = \frac{2E\left[e_a^2(k)\right]}{E\left[f(e(k))e^2(k)\right]\mathrm{Tr}\left[\mathbf{R}_{\mathbf{uu}}\right]} \tag{22}$$

where $\mathrm{Tr}(\cdot)$ denotes the trace of a matrix. In particular, if the variance of the background noise σ_η^2 is much smaller than $E\left[e_a^2(k)\right]$, i.e., $\sigma_\eta^2 \ll E\left[e_a^2(k)\right]$, we have $E\left[e^2(k)\right] = E\left[e_a^2(k)\right] + \sigma_\eta^2 \approx E\left[e_a^2(k)\right]$. Recalling Assumption **A2**, (22) is simplified as

$$\mu_m = \frac{2}{E\left[f(e(k))\right]\mathrm{Tr}\left[\mathbf{R}_{\mathbf{uu}}\right]} \tag{23}$$

From Fig. 2, the nonlinear function $f(e(k))$ is an even function. Moreover, when $e(k) > 0$, the function $f(e(k))$ is monotonically decreasing. Note that in (14), a lower bound is set for $|e(k)|$, i.e., $\left[|e(k)|\right]_{\min} = \kappa\sigma_\eta$. Therefore, $f(e(k))$ takes the maximum value at point $\kappa\sigma_\eta$. The minimum upper bound $\bar{\mu}_m$ of the step size is given by

$$\bar{\mu}_m = \frac{2}{f(\kappa\sigma_\eta)\mathrm{Tr}\left[\mathbf{R}_{\mathbf{uu}}\right]} \tag{24}$$

In the case of $\sigma_\eta^2 \ll E\left[e_a^2(k)\right]$, a sufficient condition for the mean-square convergence can be expressed as

$$0 < \mu < \frac{2}{f(\kappa\sigma_\eta)\mathrm{Tr}\left[\mathbf{R}_{\mathbf{uu}}\right]} \tag{25}$$

IV. Simulation Results

In this section, the performance of the proposed algorithm is evaluated in the system identification scenario. Experiments are performed in an impulsive noise environment which introduces both the background noise and impulsive noise. Background noise is a white Gaussian process resulting in a signal-to-noise ratio (SNR) of 15dB. The unknown system is given by $\mathbf{w}_o = 0.5 \times \mathrm{ones}(1, M)$ with $M = 11$. The zero-mean, white Gaussian signal with unit variance is used as the input. The normalized mean-square-deviation (NMSD), defined as $20\log_{10}\left[\|\mathbf{w}_o - \mathbf{w}(k)\|^2 / \|\mathbf{w}_o\|^2\right]$, is employed to evaluate the algorithm performance. Generally, the impulsive noise can be modeled by either the Bernoulli-Gaussian (BG) distribution [9], [17] or the α -Stable distribution [10], [18]. We consider both cases. In the simulations, when performing the comparison of various algorithms, we let the unknown vector $\mathbf{w}_o$ change to $-\mathbf{w}_o$ at the middle of iterations to evaluate their tracking abilities. All results are the average of 100 independent trials.

A. *Bernoulli-Gaussian Distribution*

The impulsive noise is modeled by a Bernoulli-Gaussian process, i.e. $v(k) = q(k)h(k)$, where $q(k)$ is a white Gaussian process with variance $\sigma_q^2 = 100E\left[\left(\mathbf{u}^T(k)\mathbf{w}_o\right)^2\right]$, and $h(k)$ is a Bernoulli process with the probability mass function given by $P\left[h(k)=1\right] = P_r$ and $P\left[h(k)=0\right] = 1 - P_r$ (P_r denotes the probability of the occurrence of impulsive interference). In this example, we set $P_r = 0.01$.

Fig. 3 shows the NMSD curves of the proposed GVSS-CMPN algorithm for different κ. The regulating factor is fixed as $\theta = -2$. As can be seen, when $\kappa = 1$, the steady-state misalignment of the GVSS-CMPN algorithm is higher than that when $\kappa < 1$. Moreover, the GVSS-CMPN algorithm with $\kappa = 0.1, 0.01, 0.001$ achieves almost the same performance. Therefore, in our simulation, we set $\kappa = 0.1$.

Fig. 4 shows the NMSD curves of the proposed GVSS-CMPN algorithm for different θ. As can be seen, the GVSS-CMPN algorithm with $\theta = 2$ exhibits faster convergence than that with other values, but increases its steady-state misalignment. With the decrease of θ, the proposed algorithm is gradually decelerating the convergence meanwhile reducing the steady-state misalignment. To ensure good robustness, for the GVSS-CMPN algorithm, we set $\theta = -2$.

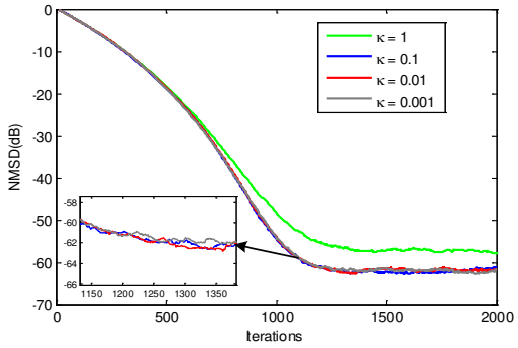

Fig. 3. NMSD performance of the proposed GVSS-CMPN algorithm against κ; $\mu = 0.002$.

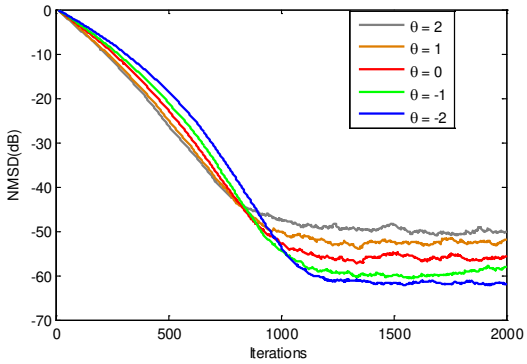

Fig. 4. NMSD performance of the proposed GVSS-CMPN algorithm against θ; $\mu = 0.002$.

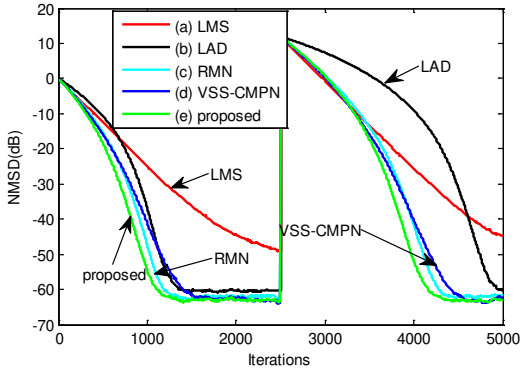

Fig. 5. NMSD performance of various algorithms. (a) $\mu = 0.002$ (b) $\mu = 0.002$ (c) $\mu = 0.002$ (d) $\mu = 0.0015$ (e) $\mu = 0.002$.

Fig. 5 compares the NMSD curves of the LMS, LAD, RMN, VSS-CMPN and GVSS-CMPN algorithms. For the RMN algorithm, we select $N_w = 10$ as suggested in [7]. As can be

seen from Fig. 5, the LMS algorithm performs poorly since it cannot combat the impulsive noise. In contrast, the LAD, RMN and VSS-CMPN algorithms have great improvement for both the convergence rate and steady-state misalignment. Finally, the GVSS-CMPN algorithm provides superior performance compared to other algorithms.

B. α -Stable Distribution

The impulsive noise is here described by the α -stable distribution with a characteristic function $\varphi(x) = \exp(-\gamma|x|^{\alpha})$, where the characteristic exponent $\alpha \in (0, 2]$ describes the impulsiveness of the noise (smaller α leads to more impulsive noise samples) and $\gamma > 0$ characterizes the dispersion level of the noise [10], [18]. In particular, when $\alpha = 2$, it degenerates to the Gaussian noise. In this example, we consider two cases of α and set $\gamma = 1/15$. For the proposed algorithm, we set $\kappa = 0.01$ and $\theta = -2$.

As can be seen from Figs. 6 and 7, the proposed GVSS-CMPN algorithm behaves better than other tested algorithms as well as retains good robustness.

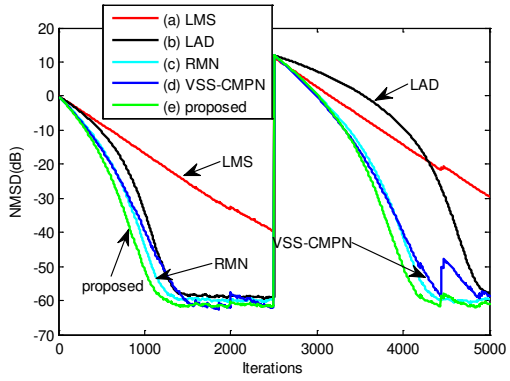

Fig. 6. NMSD performance of various algorithms, and $\alpha = 1.3$ (a) $\mu = 0.001$ (b) $\mu = 0.002$ (c) $\mu = 0.002$, $N_w = 10$ (d) $\mu = 0.0015$ (e) $\mu = 0.002$.

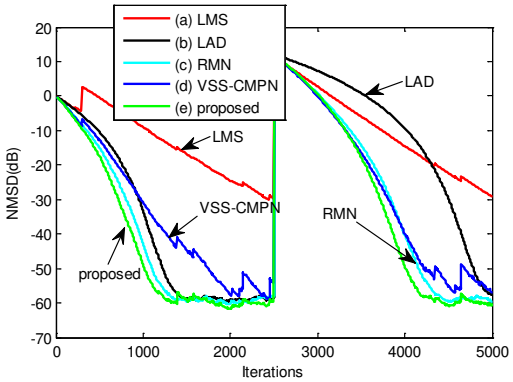

Fig. 7. NMSD performance of various algorithms, and $\alpha = 1.2$. (a) $\mu = 0.001$ (b) $\mu = 0.002$ (c) $\mu = 0.002$, $N_w = 10$ (d) $\mu = 0.0015$ (e) $\mu = 0.002$

V. Conclusion

In this brief, we have designed a linear function for the probability density-like function to derive the GVSS-CMPN algorithm. The existing VSS-CMPN algorithm can be viewed as a special case of the proposed GVSS-CMPN algorithm. We have discussed the influence of the regulating factor based on a nonlinear function of the error. In addition, we have presented the mean-square stability analysis of the algorithm. Simulations in the context of system identification have demonstrated the advantages of the proposed algorithm over known techniques. In our future work, we will investigate how the algorithm performance can be improved by adapting the regulating factor to the signal statistics.

References

[1] A. H. Sayed, *Fundamentals of Adaptive Filtering*. New York: Wiley, 2003.

[2] F. Huang, J. Zhang, and S. Zhang, "Combined-step-size affine projection sign algorithm for robust adaptive filtering in impulsive interference environments," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 63, no. 5, pp. 493–497, May 2016.

[3] S. Zhang, J. Zhang, and H. Han, "Robust shrinkage normalized sign algorithm in an impulsive noise environment," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 64, no. 1, pp. 91–95, Jan. 2017.

[4] F. Huang, J. Zhang, and S. Zhang, "Combined-step-size normalized subband adap- tive filter with a variable-parametric step-size scaler against impulsive interferences," *IEEE Trans. Circuits Syst. II, Exp. Briefs* (2017), doi: 10.1109/TCSII. 2017.2771430.

[5] V. J. Mathews and S. H. Cho, "Improved convergence analysis of stochastic gradient adaptive filters using the sign algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, no. 4, pp. 450–454, Apr. 1987

[6] J. A. Chambers, O. Tanrikulu, and A. G. Constantinides, "Least mean mixed-norm adaptive filtering," *Electron. Lett*., vol. 30, pp. 1574–1575, Sept. 1994

[7] J. A. Chambers and A. Avlonitis, "A robust mixed-norm adaptive filter algorithm," *IEEE Signal Process. Lett*., vol. 4, no. 4, pp. 46–48, Feb. 1997.

[8] E. V. Papoulis and T. Stathaki, "A normalized robust mixed-norm adaptive algorithm for system identification," *IEEE Signal Process. Lett*., vol. 11, no. 1, pp. 56–59, Jan. 2004.

[9] H. Zayyani, "Continuous mixed p-norm adaptive algorithm for system identification," *IEEE Signal Process. Lett*., vol. 21, no. 9, pp. 1108–1110, Sept. 2014.

[10] L. Lu, H. Zhao, and B. Chen, "Robust adaptive algorithm for smart antenna system with α-stable noise," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, DOI:10.1109/TCSII.2017.2761358, 2017

[11] L. Lu and H. Zhao, "Adaptive Volterra filter with continuous l_p-norm using a logarithmic cost for nonlinear active noise control," *J. Sound Vib*., vol. 364, pp. 14–29, Mar. 2016.

[12] W. Ma, B. Chen, H. Qu, and J. Zhao, "Sparse least mean p-power algorithms for channel estimation in the presence of impulsive noise," *Signal Image Video Process*., vol. 10, no. 3, pp. 503–510, Mar. 2016

[13] B. Chen, L. Xing, Z. Wu, J. Liang, J. C. Pr´ıncipe, and N. Zheng, "Smoothed least mean p-power error criterion for adaptive filtering," *Digit. Signal Process*., vol. 40, no. 1, pp. 154–163, 2015

[14] C. Paleologu, S. Ciochină, and J. Benesty, "Variable step-size NLMS algorithm for under-modeling acoustic echo cancellation," *IEEE Signal Process. Lett*., vol. 15, pp. 5–8, 2008.

[15] H.-S. Lee, S.-H. Yim, and W.-J. Song, "z^2-proportionate diffusion LMS algorithm with mean square performance analysis," *Signal Processing*., vol. 131, pp. 154–160, Feb. 2017.

[16] S. Wang, L. Dang, B. Chen, S. Duan, L. Wang, and C. K. Tse, "Random Fourier filters under maximum correntropy criterion," *IEEE Transactions on Circuits and Systems I: Regular Papers*., DOI: 10.1109/TCSI.2018.2825241, 2018.

[17] T. Shao, Y. R. Zheng, and J. Benesty, "An affine projection sign algorithm robust against impulsive interferences," *IEEE Signal Process. Lett*., vol. 17, no. 4, pp. 327–330, Apr. 2010.

[18] K. Pelekanakis and M. Chitre, "Adaptive sparse channel estimation under symmetric alpha-stable noise," *IEEE Transactions on Wireless Communications*., vol. 13, no. 6, pp. 3183–3195, Jun. 2014